DOCKET NO. 03-SIN-092 (STMI01-03092) **PATENT**

APPARATUS AND METHOD FOR RENDERING AUDIO INFORMATION TO VIRTUALIZE SPEAKERS IN AN AUDIO SYSTEM

Inventors:

Yuan Wu
Block 690, #12-209, Jurong West Central 1
Singapore 640690
Citizen of China

Sapna George
42, Li Hwan Drive
Singapore 557076
Citizen of India

Assignee:

STMicroelectronics Asia Pacific Pte., Ltd.
20 Science Park Road #01-28/30
Teletech Park
Singapore 117674

William A. Munck
John T. Mockler
DAVIS MUNCK, P.C.
P.O. Drawer 800889
Dallas, Texas 75360
(972) 628-3600

APPARATUS AND METHOD FOR RENDERING AUDIO INFORMATION TO VIRTUALIZE SPEAKERS IN AN AUDIO SYSTEM

TECHNICAL FIELD

[0001] This disclosure is generally directed to sound processing systems and more specifically to an apparatus and method for rendering audio information to virtualize speakers in an audio system.

BACKGROUND

[0002] Multi-channel sound systems have become increasingly popular in recent years. While older sound systems often included two speakers placed in front of a listener, multi-channel systems typically use more than two speakers. As an example, in a 5.1 audio system, five speakers and a subwoofer are placed around the listener. In this type of audio system, one speaker is typically placed directly in front of the listener, two speakers in front and to the sides of the listener, and two speakers to the sides and possibly behind the listener. These multi-channel systems typically produce more realistic sound effects, such as more realistic surround sound playback during a movie.

[0003] Despite the popularity of these multi-channel systems, many people continue to use conventional two-speaker systems. The use of two speakers in an audio system typically limits or prevents the audio system from producing more realistic sounds using the speakers.

SUMMARY

[0004] This disclosure provides an apparatus and method for rendering audio information to virtualize speakers in an audio system.

[0005] In one aspect, an audio processor includes a virtualizer operable to process audio information to virtualize at least one speaker so that, from a listener's perspective, sounds appear to come from at least one direction where a physical speaker is not present. The audio processor also includes a controller operable to configure the virtualizer. The virtualizer can be configured to virtualize the at least one speaker at any location in a space around the listener.

[0006] In another aspect, a method includes generating first output signals for a first physical speaker and generating second output signals for a second physical speaker. The first output signals emulate effects of a virtual speaker on one ear of a listener, and the second output signals emulate effects of the virtual speaker on another ear of the listener. Each of the output signals also at least partially cancels crosstalk caused by the other output signals.

[0007] One or more technical features may be present according to various embodiments of this disclosure. Particular embodiments of this disclosure may exhibit none, some, or all of the following

features depending on the implementation. In one embodiment, a system for rendering audio information to virtualize speakers is provided. In particular, the system is capable of rendering audio information so that, from the perspective of a listener, sounds appear to come from one or more directions where speakers are not present. For example, the system may be capable of reproducing multi-channel sound in a two-speaker system in a more realistic fashion. In other words, using two speakers, the system makes it appear to a listener that sounds are being produced by additional "virtual" speakers around the listener.

[0008] In particular embodiments, the system is capable of rendering audio information for any number of virtual speakers. For example, the system could allow a two-speaker system to emulate a 5.1 audio system more realistically. In this example, the sounds produced by the two speakers may, from the listener's perspective, appear as if they were produced by five speakers around the listener.

[0009] This has outlined rather broadly several features of this disclosure so that those skilled in the art may better understand the DETAILED DESCRIPTION that follows. Additional features may be described later in this document. Those skilled in the art should appreciate that they may readily use the concepts and the specific embodiments disclosed as a basis for modifying or designing other

structures for carrying out the same purposes of this disclosure. Those skilled in the art should also realize that such equivalent constructions do not depart from the spirit and scope of the invention in its broadest form.

[0010] Before undertaking the DETAILED DESCRIPTION below, it may be advantageous to set forth definitions of certain words and phrases used throughout this patent document. The terms "include" and "comprise," as well as derivatives thereof, mean inclusion without limitation. The term "or" is inclusive, meaning and/or. The phrases "associated with" and "associated therewith," as well as derivatives thereof, may mean to include, be included within, interconnect with, contain, be contained within, connect to or with, couple to or with, be communicable with, cooperate with, interleave, juxtapose, be proximate to, be bound to or with, have, have a property of, or the like. The term "controller" means any device, system, or part thereof that controls at least one operation. A controller may be implemented in hardware, firmware, or software, or a combination of at least two of the same. It should be noted that the functionality associated with any particular controller may be centralized or distributed, whether locally or remotely. Definitions for certain words and phrases are provided throughout this patent document, and those of ordinary skill in the art should understand that in many, if not most

instances, such definitions apply to prior as well as future uses of such defined words and phrases.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] For a more complete understanding of this disclosure and its features, reference is now made to the following description taken in conjunction with the accompanying drawings, in which:

[0012] FIGURE 1 illustrates an example audio system according to one embodiment of this disclosure;

[0013] FIGURES 2A and 2B illustrate example audio/video devices according to one embodiment of this disclosure;

[0014] FIGURE 3 illustrates an example virtualization of a speaker according to one embodiment of this disclosure;

[0015] FIGURE 4 illustrates an example audio virtualizer for virtualizing one speaker according to one embodiment of this disclosure;

[0016] FIGURE 5 illustrates an example audio virtualizer for virtualizing two speakers according to one embodiment of this disclosure;

[0017] FIGURE 6 illustrates an example audio virtualizer for virtualizing *n* speakers according to one embodiment of this disclosure;

[0018] FIGURES 7A and 7B illustrate an example audio virtualizer for emulating a 5.1 audio system according to one embodiment of this disclosure;

[0019] FIGURES 8A through 8C illustrate another example audio

virtualizer for emulating a 5.1 audio system according to one embodiment of this disclosure; and

[0020] FIGURE 9 illustrates an example method for rendering audio information to virtualize speakers according to one embodiment of this disclosure.

DETAILED DESCRIPTION

[0021] FIGURE 1 illustrates an example audio system 100 according to one embodiment of this disclosure. In the illustrated example, the audio system 100 includes an audio/video device 102 and two speakers 104a and 104b. Other embodiments of the audio system 100 may be used without departing from the scope of this disclosure.

[0022] The audio/video device 102 is coupled to the speakers 104a and 104b. The audio/video device 102 could also be coupled to a subwoofer 106. In this document, the term "couple" and its derivatives refer to any direct or indirect communication between two or more elements, whether or not those elements are in physical contact with one another. The audio/video device 102 receives or generates audio information, which is sent to the speakers 104 and possibly the subwoofer 106 for presentation to one or more listeners 108. In this document, the phrase "audio information" refers to any signal, pattern, or other information that symbolizes, characterizes, or otherwise represents audio sounds, whether the information is in digital, analog, or other form.

[0023] The audio/video device 102 represents any device, system, or part thereof that is capable of providing audio information to one or more speakers 104. The audio/video device 102 could also include functionality for receiving or generating video information

for display on a television 110 or other display device. As particular examples, the audio/video device 102 could represent a television tuner or receiver, a compact disk ("CD") player, a digital versatile disk ("DVD") player, an audio tuner or receiver, a desktop, laptop, or server computer, or any other suitable device.

[0024] In one aspect of operation, the audio/video device 102 is capable of rendering audio information to create the appearance of one or more "virtual" speakers 112a-112e. A virtual speaker 112 represents a direction from which the listener 108 believes sounds are originating. In other words, the two actual speakers 104 produce sounds that the listener 108 believes are coming from one or more directions other than from the speakers 104. For example, the audio/video device 102 could make it appear as if sounds are coming from a center speaker 112a directly in front of the listener 108. The audio/video device 102 could also make it appear as if sounds are coming from two surround sound speakers 112b and 112c to the sides of and possibly behind the listener 108. In addition, the audio/video device 102 could make it appear as if sounds are coming from two front speakers 112d and 112e in front of and to the sides of the listener 108.

[0025] The audio/video device 102 includes any hardware, software, firmware, or combination thereof for virtualizing one or

more speakers 112. Example embodiments of the audio/video device 102 are shown in FIGURES 2A and 2B, which are described below. Although FIGURE 1 has described the audio system 100 as including an audio/video device 102, a device 102 that omits the video functionality could also be used in the audio system 100.

[0026] While FIGURE 1 has shown two physical speakers 104 virtualizing one or more virtual speakers 112, the system 100 could include any number of physical speakers 104. Also, any number of physical speakers 104 could be used to virtualize at least one virtual speaker 112. For example, three speakers 104 could be used in the system 100, and two of the three speakers 104 could be used to virtualize two additional virtual speakers 112. As a particular example, a system 100 could include three speakers 104 (two as shown in FIGURE 1, one in the position of a virtual speaker 112), and the two speakers 104 in front of the listener 108 could virtualize the remaining four virtual speakers 112 shown in FIGURE 1.

[0027] Although FIGURE 1 illustrates one example of an audio system 100, various changes may be made to FIGURE 1. For example, the number and positions of the virtual speakers 112 shown in FIGURE 1 are for illustration only. The audio/video device 102 could virtualize any number of speakers 112 at any location or locations without departing from the scope of this disclosure.

Also, the audio system 100 could include any number of real speakers 104.

[0028] FIGURES 2A and 2B illustrate example audio/video devices 102 according to one embodiment of this disclosure. In these example embodiments, the audio/video device 102 includes an audio/video source 202, an audio processor 204, a memory 206, and two outputs 208. Other embodiments of the audio/video devices 102 could be used without departing from the scope of this disclosure.

[0029] In FIGURE 2A, the audio/video source 202 is coupled to the audio processor 204. The audio/video source 202 represents any suitable source of audio information. For example, the audio/video source 202 could represent a CD or DVD reader capable of extracting audio information from a CD or DVD. The audio/video source 202 could also represent a radio tuner capable of capturing transmitted radio signals. The audio/video source 202 could further represent a television tuner, such as a high definition television ("HDTV") tuner, capable of capturing transmitted television signals that include audio signals. The audio/video source 202 could represent any other or additional source of audio information.

[0030] Audio information from the audio/video source 202 is provided to the audio processor 204. The audio processor 204 processes the audio information for presentation to one or more listeners 108. For example, the audio processor 204 could process

the audio information to virtualize one or more virtual speakers 112. The audio processor 204 includes any hardware, software, firmware, or combination thereof for processing audio information. As particular examples, the audio processor 204 could include one or more microprocessors, digital signal processors ("DSPs"), field programmable gate arrays ("FPGAs"), application specific integrated circuits ("ASICs"), or any other suitable processor or processors.

[0031] In the illustrated example, the audio processor 204 includes a virtualizer 210 and a controller 212. The virtualizer 210 and controller 212 could, for example, represent different hardware components or different software programs executed by the audio processor 204.

[0032] The virtualizer 210 receives the audio information from the audio/video source 202 and processes the audio information to virtualize one or more speakers 112. For example, the virtualizer 210 could process the audio information to virtualize a speaker 112a directly in front of the listener 108. The virtualizer 210 could also process the audio information to virtualize two surround sound speakers 112b and 112c to the sides of the listener 108.

[0033] The virtualizer 210 virtualizes one or more speakers 112 based on the psycho-acoustical properties of the human auditory system. When sound waves reach a person, the person's eardrums respond to the sound waves, and the brain analyzes the responses of

both eardrums. Based on this analysis, the brain makes a judgment about the location where the sound waves originated.

[0034] In some embodiments, the response of an eardrum to sound sources at certain locations in space can be described using the concepts of Head-Related Impulse Responses ("HRIP") and Head-Related Transfer Functions ("HRTF"). A Head-Related Impulse Response is defined as the response of an eardrum excited by an impulse signal from a certain point in space. The HRIP is typically a function of azimuth, elevation, and range in relation to the source of an impulse signal. In particular embodiments, for "far field" situations where the range exceeds a threshold (such as one meter), the HRIP may be considered invariable to range.

[0035] The Head-Related Transfer Function is defined as the frequency response of the eardrum towards a certain point in space. The HRTP represents the Fourier transform of the HRIP. For far field situations, at an elevation of zero degree, the HRTF is a function of azimuth θ and can be denoted as $H(\theta)$. Measured HRTFs with different experimental conditions are available, such as in the CIPIC Interface Laboratory's CIPIC HRTF database and MIT Media lab's HRTF Measurements of a KEMAR Dummy-Head Microphone.

[0036] If a speaker was physically present at the location of a virtual speaker 112, impulse responses would be received at the ears of the listener 108. To create a virtual speaker 112, the

ears of the listener 108 should receive the same or similar impulse responses from the actual speakers 104 that would be received if a real speaker was present at the location of the virtual speaker 112. In some embodiments, the virtualizer 210 makes use of the characteristics of HRTFs during the virtualization process. The virtualizer 210 includes any hardware, software, firmware, or combination thereof for virtualizing one or more speakers 112. Example virtualizers 210 are shown in FIGURES 4-6, 7A, and 8A, and the operation of these virtualizers 210 are described below.

[0037] The controller 212 controls the operation of the virtualizer 210. For example, in some embodiments, the virtualization of the speakers 112 can be customized based on parameters 214-218 stored in the memory 206. The controller 212 represents any hardware, software, firmware, or combination thereof for configuring or otherwise controlling the operation of the virtualizer 210.

[0038] The memory 206 is coupled to the audio processor 204. The memory 206 stores and facilitates retrieval of information used by the audio processor 204 to process audio information. For example, the memory 206 may store the parameters 214-218 used by the controller 212 to configure the virtualizer 210. The memory 206 includes any hardware, software, firmware, or combination thereof for storing and facilitating retrieval of information, such

as a volatile or non-volatile device or devices.

[0039] The memory 206 stores and the controller 212 uses any suitable parameters to configure the virtualizer 210. For example, as described above, the virtualizer 210 may use HRTFs to virtualize one or more speakers 112. HRTFs typically vary based on individual listeners 108 and on the position of the actual speakers 104. Also, different listeners 108 often have different preferences about the locations of the virtual speakers 112. In this example, the virtualization of the speakers 112 can be based on parameters such as the position 214 of the actual speakers 104, the number or location 216 of the virtual speaker or speakers 112, and information about the HRTFs 218 of a listener 108. Other or additional parameters could also be used by the controller 212. The controller 212 collects these parameters and configures the virtualizer 210 to give the desired audio effect.

[0040] The audio information processed by the audio processor 204 is provided to the two speakers 104 through outputs 208a and 208b. The outputs 208 represent any suitable structure or device capable of providing audio information to the speakers 104. For example, the outputs 208 could represent connectors capable of accepting RCA-type cables or two-wire speaker cables.

[0041] Although FIGURE 2A illustrates an audio/video source 202 in an audio/video device 102, the device 102 could represent an

audio-only device. In these embodiments, the audio device 102 could use an audio source 202 that does not provide any video information. When video information is provided by the audio/video source 202, the video information is sent to a video processor 220. The video processor 220 processes the video information for display on a television 110 or other display device. For example, the video processor 220 may process the video information so that it can be displayed on a Red/Green/Blue ("RGB") device, a Video Graphics Array ("VGA") device, an HDTV device, or a plasma display. The processed video information may be provided to the display device through one or more outputs 222, such as a digital coaxial output or component video outputs.

[0042] FIGURE 2B illustrates another example embodiment of an audio/video device 102. In this example, the audio/video device 102 is similar to the device 102 shown in FIGURE 2A. In addition to the components described above with respect to FIGURE 2A, the audio/video device 102 in FIGURE 2B includes an audio decoder 250. In this example embodiment, the audio/video source 202 provides audio information that has been encoded, such as audio information that has been encoded using the 5.1 or other multi-channel standard. The audio decoder 250 receives and decodes the encoded audio information. In decoding the audio information, the audio decoder 250 may separate the audio information into the various

channels 252a-252e. As a particular example, the audio decoder 250 may separate the audio information into left and right front channels 252a and 252b, left and right surround sound channels 252c and 252d, and a center channel 252e. Other decoding schemes associated with any number of channels may be used by the audio decoder 250. The audio decoder 250 includes any hardware, software, firmware, or combination thereof for decoding audio information.

[0043] In this example embodiment, the controller 212 in the audio processor 204 also uses a listening mode parameter 254 to configure the virtualizer 210. In some embodiments, the audio processor 204 can virtualize the location of the speakers 112 differently to alter the perceived position of one or more of the virtual speakers 112. The different perceived positions of the virtual speakers 112 may correspond to different listening modes that can be selected by a listener 108. As a particular example, the virtual surround sound speakers 112b and 112c could be located either directly to the sides of the listener 108 or to the sides and behind the listener 108, depending on the listening mode parameter 254 selected. As another example, the virtual front speakers 112d and 112e may or may not be virtualized, depending on the listening mode parameter 254 selected. Based on the listening mode parameter 254, the controller 212 decides which channels

should be virtualized, and the virtualizer 210 processes the audio signals according to the decisions made by the controller 212.

[0044] Although FIGURES 2A and 2B illustrate example embodiments of an audio/video device 102, various changes may be made to FIGURES 2A and 2B. For example, the video processor 220 need not be provided in the devices 102. Also, FIGURES 2A and 2B have been simplified for ease of illustration and explanation. Other embodiments of the devices 102 including other or additional components could also be used. In addition, the functional divisions shown in FIGURES 2A and 2B are for illustration only. Various components could be combined or omitted and additional components could be added according to particular needs.

[0045] FIGURE 3 illustrates an example virtualization 300 of a virtual speaker 112 according to one embodiment of this disclosure. In particular, FIGURE 3 illustrates the virtualization of a virtual surround sound speaker 112b that is positioned to the left and behind a listener 108. Although FIGURE 3 describes the virtualization of this particular virtual speaker 112b in a particular location, the principles shown and described below can be used to virtualize one or multiple speakers 112 at any suitable location or locations.

[0046] As described above, in some embodiments, the virtualizer 210 uses HRTFs to virtualize one or more virtual speakers 112. The

example shown in FIGURE 3 illustrates the creation of a virtual speaker 112b that is closer to the left ear of the listener 108. In this example, for ease of illustration and explanation, the space around the listener 108 is divided into two halves by a centerline 302. Also shown in FIGURE 3 is the angle (θ) 304 between the centerline 302 and each physical speaker 104 and the angle (ϕ) 306 between the centerline 302 and the virtual speaker 112b.

[0047] If a speaker was physically present at the illustrated location of the virtual speaker 112, the left ear of the listener 108 would first receive sound waves from the speaker 112b. After some amount of time, the right ear of the listener 108 would receive sound waves from the speaker 112b. The transfer function from the virtual speaker 112b to the listener's left ear is represented as $H_i(\phi)$. The transfer function from the virtual speaker 112b to the listener's right ear is represented as $H_c(\phi)$. The time difference, $t(\phi)$, between the sound waves from the speaker 112b arriving at the listener's ears is defined as the inter-time difference (ITD). Similarly, the transfer function from the left speaker 104a to the listener's left ear is represented as $H_i(\theta)$, and the transfer function from the left speaker 104a to the listener's right ear is represented as $H_c(\theta)$. The inter-time

difference between the sound waves from the speaker 104a arriving at the listener's ears is represented as $t(\theta)$.

[0048] To create the appearance of a virtual speaker 112b, the left speaker 104a emulates the impact of the virtual speaker's sound waves on the listener's left ear. The right speaker 104b emulates the impact of the virtual speaker's sound waves on the listener's right ear. To emulate the impact to the listener's left ear, the sounds S to be produced by the left speaker 104a are transformed by $\frac{H_i(\phi)}{H_i(\theta)}$. Similarly, to emulate the impact to the listener's right ear, the sounds produced by the right speaker 104b are transformed by $\frac{H_c(\phi)}{H_i(\theta)}$, which is also equal to $S \times \frac{H_i(\phi)}{H_i(\theta)} \times \frac{H_c(\phi)}{H_i(\phi)} = S_i \times \frac{H_c(\phi)}{H_i(\phi)}$, where S_i represents the original audio signal S after being filtered by $\frac{H_i(\phi)}{H_i(\theta)}$.

[0049] Ideally, based on these properties, the virtualizer 210 could produce S_i by filtering the original signal S with a filter having a response of $\frac{H_i(\phi)}{H_i(\theta)}$ and sending the resulting signal to the left speaker 104a. The virtualizer 210 could also filter S_i using a filter with a response of $\frac{H_c(\phi)}{H_i(\phi)}$ and send the resulting signal to

the right speaker 104b. These signals would ideally emulate the virtual speaker 112b.

[0050] As shown in FIGURE 3, however, the left speaker 104a has an impact on the right ear of the listener 108, and the right speaker 104b has an impact on the left ear of the listener 108. The effect that a speaker 104 has on the opposite ear of the listener 108 is referred to as "crosstalk." Crosstalk interferes with the ideal operation of the speakers 104, meaning that it can interfere with or destroy the effect of the virtualization. As described below, to reduce or eliminate crosstalk, the output of each speaker 104 is used to generate an out-of-phase signal for the other speaker 104. The out-of-phase signals help to reduce or cancel the crosstalk produced by the speakers 104, which helps to more effectively virtualize the speaker 112b.

[0051] Although FIGURE 3 illustrates one example of the virtualization 300 of a virtual speaker 112b, various changes may be made to FIGURE 3. For example, any other or additional virtual speakers 112 could be emulated by the speakers 104. Also, the speakers 104 could have any position with respect to the listener 108. As an example, while FIGURE 3 illustrates that each speaker 104 is positioned at the same angle 304 from the centerline 302, each speaker 104 could be placed at different angles 304 from the centerline 302.

[0052] FIGURE 4 illustrates an example audio virtualizer 210 for virtualizing one speaker 112 according to one embodiment of this disclosure. In the illustrated example, the speaker 112 being virtualized is closer to the left ear of the listener 108. The same or similar structure could be used to virtualize a speaker 112 closer to the right ear of the listener 108.

[0053] As described above, the sounds produced by a real speaker at the location of a virtual speaker 112b would have a transfer function of $H_i(\phi)$ for the listener's left ear, a transfer function of $H_c(\phi)$ for the listener's right ear, and an inter-time difference $t(\phi)$. Based on this, the virtualizer 210 in FIGURE 4 receives an input signal 402 and processes the input signal 402 so that the speakers 104 produce sounds with the proper transfer functions and inter-time difference.

[0054] In this example, the input signal 402 for the left speaker 104a is provided to a filter 404. The response of the filter 404, P_L, may be determined using the formula:

$$|P_L| = \left|\frac{H_i(\phi)}{H_i(\theta)}\right|.$$

This transform alters the input signal 402 to produce a filtered input signal 406. In the absence of crosstalk, the filtered signal 406 would be provided to the left speaker 104a, and it would allow the left speaker 104a to emulate the effects of the virtual speaker

112 on the listener's left ear.

[0055] The filtered signal 406 is also provided to a forward crossover path 408. The forward crossover path 408 processes the filtered signal 406 before providing it to the right speaker 104b. In this example, the forward crossover path 408 includes a filter 410 and a delay line 412.

[0056] Ideally, HRTFs contain the proper inter-time difference, and the virtualizer 210 need not alter or provide an extra delay to the signals to emulate the inter-time difference. However, this may require unstable filters having high orders, which are inefficient. Simpler filters and delay lines can be used to approximate the needed filter response.

[0057] The filter 410 receives the signal 406 produced by the filter 404 and further filters the signal 406 to produce a signal 414. The response of the filter 410, F_L, may be determined using the formula:

$$|F_L| = \left|\frac{H_c(\phi)}{H_i(\phi)}\right|.$$

In the absence of crosstalk, the signal 414 would be provided to the right speaker 104b, and it would allow the right speaker 104b to emulate the effects of the virtual speaker 112 on the listener's right ear.

[0058] Because the filter 410 approximates the filter needed to

emulate the virtual speaker 112, the filter 410 may not have the correct delay. As a result, the speakers 104 may produce sounds having an improper inter-time difference. The delay line 412 delays the signal 406 provided to the filter 410 to compensate for the inexact delay of the filter 410. The delay D_L introduced by the delay line 412 may be determined using the formula:

$$D_L = t(\phi) - t(F_L)$$

where $t(\phi)$ represents the desired inter-time difference for the virtual speaker 112, and $t(F_L)$ represents the delay introduced by the filter 410. The inter-time difference $t(\phi)$ could have any value. As an example, when the angle 306 from the centerline 302 to the virtual speaker 112 equals 90°, the inter-time difference could range from 0.65 to 0.70 ms depending on the head shape of the listener 108.

[0059] As described above, in the absence of crosstalk, the signals 406 and 414 could be used to emulate the virtual speaker 112. However, the presence of crosstalk can interfere with and possibly destroy the effective emulation of a virtual speaker 112. To compensate for crosstalk, the virtualizer 210 includes two feedback crossover paths 416a and 416b. The feedback crossover paths 416 process output signals 418, 420 provided to the two speakers 104. Each feedback crossover path 416 takes the output to

one speaker 104 and generates an out-of-phase signal 422 for the other speaker 104. The out-of-phase signal 422 allows one speaker 104 to cancel the crosstalk produced by the other speaker 104.

[0060] In the illustrated example, each feedback crossover path 416 includes a filter 424 and a delay line 426. The filter 424 receives one of the output signals 418, 420 and filters the output signal to produce the out-of-phase signal 422. The response of the filter 424, F_T, may be determined using the formula:

$$|F_T| = \left|\frac{H_c(\theta)}{H_i(\theta)}\right|.$$

[0061] Because the filter 424 may approximate the needed filter response, the filter 424 may have an incorrect delay. The delay line 426 delays the output signal 418, 420 provided to the filter 424 to compensate for the inexact delay of the filter 424. The delay D_T introduced by the delay line 426 may be determined using the formula:

$$D_T = t(\theta) - t(F_T)$$

where $t(\theta)$ represents the inter-time difference for left speaker 104, and $t(F_T)$ represents the delay introduced by the filter 424.

[0062] The output signals 418, 420 provided to the speakers 104 represent combinations of the various signals produced by the filter 404, the forward crossover path 408, and the feedback crossover paths 416. For example, a combiner 428 produces the

output signal 418 for the left speaker 104a by combining the signal 406 produced by the filter 404 and the out-of-phase signal 422a produced by the feedback crossover path 416a. In this way, the left speaker 104a uses the output signal 418 to emulate the effects of the virtual speaker 112 on the left ear of the listener 108 while canceling crosstalk from the right speaker 104b. A combiner 430 produces the output signal 420 for the right speaker 104b by combining the signal 414 produced by the forward crossover path 408 and the out-of-phase signal 422b produced by the feedback crossover path 416b. In this way, the right speaker 104b uses the output signal 420 to emulate the effects of the virtual speaker 112 on the right ear of the listener 108 while canceling crosstalk from the left speaker 104a.

[0063] The HRTFs and inter-time difference used by the virtualizer 210 can vary from listener 108 to listener 108. For example, they may vary based on the positions of the speakers 104 and the body shape and dimensions of the listener 108. The placement of speakers 104 (defined by the angle 304) affects $H_i(\theta)$, $H_c(\theta)$, and $t(\theta)$. The location of the virtual speaker 112 (defined by angle 306) affects $H_i(\phi)$, $H_c(\phi)$, and $t(\phi)$. The virtualizer 210 can be configured by the controller 212 to take the various parameters into account when virtualizing a speaker 112. In particular, the virtualizer 210 can be configured by altering the

responses of the filters 404, 410, 424 and the delay lines 412, 426 accordingly. The virtualizer 210 could also be configured in a non-individualized manner, such as by assuming default values for the angles 304 and 306.

[0064] Each of the filters 404, 410, 424 in FIGURE 4 could represent any hardware, software, firmware, or combination thereof for filtering signals. As particular examples, the filters 404, 410, 424 could represent Finite Impulse Response ("FIR") or Infinite Impulse Response ("IIR") filters. Each of the delay lines 412, 426 could represent any hardware, software, firmware, or combination thereof for delaying a signal. As a particular example, the delay lines 412, 426 may be implemented as circular buffers. In addition, as shown in FIGURE 4, the out-of-phase signal 422 produced by each feedback crossover path 416 is inverted (subtracted). In some embodiments, the inversion of the out-of-phase signals 422 can be integrated into and performed by the filters 424. This may be done, for example, when the virtualizer 210 is implemented using one or more DSPs.

[0065] In particular embodiments, the amplitude of the frequency response P_L for filter 404 equals the amplitude of $\frac{H_i(\phi)}{H_i(\theta)}$, and the filter 404 has a linear phase ideally. The amplitude of the

frequency response F_L for filter 410 equals the amplitude of $\frac{H_c(\phi)}{H_i(\phi)}$, and the amplitude of the frequency response F_T for filter 424 equals the amplitude of $\frac{H_c(\theta)}{H_i(\theta)}$. The filters 410, 424 may show low-pass characteristics and, for non-individualized design, can be implemented by low-pass filters with small (first or second) orders. In addition, the filter response F_L may depend on the azimuth associated with the virtual speaker 112, and the filter response F_T may depend on the azimuth of the speakers 104.

[0066] FIGURE 5 illustrates an example audio virtualizer 210 for virtualizing two speakers according to one embodiment of this disclosure. The audio virtualizer 210 shown in FIGURE 5 virtualizes two virtual speakers 112, one closer to the listener's left ear and one closer to the listener's right ear.

[0067] The virtualizer 210 in FIGURE 5 operates in a similar manner as the virtualizer 210 shown in FIGURE 4. The virtualizer 210 in FIGURE 5 receives two input signals 502a and 502b. The input signals 502a and 502b are provided to two filters 504a and 504b, which produce two filtered signals 506a and 506b. The filtered signals 506a and 506b are provided to two forward crossover paths 508a and 508b, which process the filtered signals 506a and 506b to produce signals 514a and 514b. Each of the

forward crossover paths 508a and 508b includes a filter 510 and a delay line 512.

[0068] The virtualizer 210 in FIGURE 5 also includes two feedback crossover paths 516a and 516b. The feedback crossover paths 516 process output signals 518 and 520 that are provided to the speakers 104 and generate out-of-phase signals 522 used to cancel crosstalk. Each feedback crossover path 516 includes a filter 524 and a delay line 526.

[0069] The output signals 518, 520 provided to the speakers 104 represent combinations of the various signals produced by the filters 504, the forward crossover paths 508, and the feedback crossover paths 516. For example, a combiner 528 combines the filtered signal 506a produced by the filter 504a and the out-of-phase signal 522a produced by the feedback crossover path 516a. Another combiner 532 combines the output of the combiner 528 and the signal 514b produced by the forward crossover path 508b. The output of the combiner 532 represents the output signal 518. Similarly, a combiner 530 combines the filtered signal 506b produced by the filter 504b and the out-of-phase signal 522b produced by the feedback crossover path 516b. Another combiner 534 combines the output of the combiner 530 and the signal 514a produced by the forward crossover path 508a. The output of the combiner 534 represents the output signal 520.

[0070] The various frequency responses of the filters 504, 510, 524 and the delays introduced by the delay lines 510, 526 may be determined using the formulas provided above in FIGURE 4. The audio processor 204 simply needs to identify the various angles 304, 306 associated with the speakers 104, 112 to properly configure the filters and delay lines. Moreover, if the virtual speakers 112 are symmetrical with respect to the centerline 302, the properties of the filters and delay lines may be symmetrical.

[0071] FIGURE 6 illustrates an example audio virtualizer 210 for virtualizing *n* speakers 112 according to one embodiment of this disclosure. In this example, the *n* virtual speakers 112 are illustrated such that at least three are to the left of the centerline 302 and at least three are to the right of the centerline 302. Other positions of the virtual speakers 112 could be used.

[0072] The virtualizer 210 shown in FIGURE 6 operates in a similar manner as the virtualizers 210 shown in FIGURES 4 and 5. Each of *n* input signals 602 is provided to and filtered by one of *n* filters 604. Each of the filtered signals is then provided to one of *n* forward crossover paths 608. The virtualizer 210 also includes two feedback crossover paths 616a and 616b, each of which produces signals used to reduce or cancel crosstalk. The output signals 618 and 620 for the speakers 104 are produced by combining

various ones of the filtered signals, the signals produced by the forward crossover paths 608, and the signals produced by the feedback crossover paths 616.

[0073] The various frequency responses of the filters and the delays introduced by the delay lines may be determined using the formulas provided above in FIGURE 4. The audio processor 204 simply needs to identify the various angles 304, 306 associated with the speakers 104, 112 to properly configure the filters and delay lines. While FIGURE 6 shows at least six speakers 112 being virtualized by the audio processor 204, any number of speakers 112 could be virtualized in the same or similar manner.

[0074] FIGURES 7A and 7B illustrate an example audio virtualizer 210 for emulating a 5.1 audio system according to one embodiment of this disclosure. FIGURES 7A and 7B illustrate one example of a virtualizer 210 for emulating a 5.1 audio system. Other virtualizers 210 could also be used to emulate a 5.1 audio system.

[0075] The virtualizer 210 shown in FIGURE 7A emulates a 5.1 audio system. The 5.1 standard represents one of the dominant multi-channel audio standards currently used. In this type of audio system, one speaker 112a is typically placed directly in front of the listener 108, two speakers 112b and 112c to the sides and possibly behind the listener 108, and two speakers 112d and 112e in front and to the sides of the listener 108. While the

virtualizers 210 shown in FIGURES 4-6 have generally been described as virtualizing speakers 112 in various locations around the listener 108, the virtualizer 210 shown in FIGURE 7A virtualizes speakers 112 to emulate a specific audio standard. In particular, the front two speakers 112d and 112e in the 5.1 audio system are assumed to be located in the same positions as the actual speakers 104. The virtualizer 210 then virtualizes a center speaker 112a and two surround sound speakers 112b and 112c.

[0076] In this example, the input signals 702a and 702b for the front two speakers 112d and 112e are simply combined with other signals and output to the speakers 104. Because the front two speakers 112d and 112e are located at the same locations as the actual speakers 104, these inputs need not be further processed.

[0077] To virtualize the center speaker 112a, an attenuator 736 receives an input signal 702c for the center speaker 112a and attenuates the signal 702c by three decibels. The attenuated signal is then provided to both speakers 104. This virtualizes the center speaker 112a directly in front of the listener 108 (at an angle 306 of zero degrees).

[0078] The virtualizer 210 virtualizes the surround sound speakers 112b and 112c in the same or similar manner as shown in FIGURE 5. Input signals 702d and 702e are filtered by filters 704a and 704b, and each filtered signal is provided to a forward

crossover path 708 that includes a filter 710. The output signals 718 and 720 are fed through two feedback crossover paths 716a and 716b that each includes a filter 724. Additional output signals 718 and 720 are then produced by combining various ones of the original two input signals 702a and 702b, the attenuated input signal 702c, the filtered input signals 702d and 702e, the signals produced by the forward crossover paths 708, and the signals produced by the feedback crossover paths 716.

[0079] In particular embodiments, the amplitude of the frequency response P_S of the filters 704 may equal an approximation of the amplitude of $\frac{H_i(\phi)}{H_i(\theta)}$. For non-individualized design, the angle 304 could assume of a value of 20°, and the angle 306 could assume of a value of 100°. In this example, the filters 704 could have approximately the frequency response shown in FIGURE 7B. The filters 710 and 724 may have frequency responses with the same amplitudes as $\frac{H_c(\phi)}{H_i(\phi)}$ and $\frac{H_c(\theta)}{H_i(\theta)}$, respectively. These filters 710, 724 may both exhibit low-pass characteristics and can be approximated by low-pass filters with attenuations for non-individualized design. Assuming that the angle 306 equals 100°, a first order IIR low-pass filter with a cut-off frequency at 1500 Hz and an attenuation of 1.5 decibels can be used as the filter 710

for non-individualized design. Assuming that the angle 304 equals 20°, a first order IIR low-pass filter with a cut-off frequency at 2000 Hz and an attenuation of 4.4 decibels can be used as the filter 724.

[0080] The virtual surround sound speakers 112b and 112c can be placed in any suitable location. For conventional 5.1 audio rendering, the angle 306 from the centerline 302 for the virtual surround sound speakers 112b and 112c is typically between 90° and 120°, although any suitable angle 306 could be used. Low Frequency Effect ("LFE") signals, such as those produced by a subwoofer 106, are typically not directional and can therefore be excluded from the virtualization process. In other words, the sounds emitted by a subwoofer 106 typically have no discernable direction from the perspective of the listener 108, so there is no need to virtualize the position of the subwoofer 106.

[0081] FIGURES 8A through 8C illustrate another example audio virtualizer for emulating a 5.1 audio system according to one embodiment of this disclosure. FIGURES 8A through 8C illustrate another example of a virtualizer 210 for emulating a 5.1 audio system. Other virtualizers 210 could also be used to emulate a 5.1 audio system.

[0082] As with the virtualizer 210 shown in FIGURE 7A, the virtualizer 210 shown in FIGURE 8A emulates a 5.1 audio system.

The virtualizer 210 shown in FIGURE 8A operates according to the same principles described above with respect to the virtualizers 210 shown in FIGURES 4-6. Using these principles, the virtualizer 210 shown in FIGURE 8A virtualizes speakers 112 to emulate a specific audio standard. In this example, the front two speakers 112d and 112e in the 5.1 audio system are not located at the same locations as the actual speakers 104. The virtualizer 210 therefore virtualizes a center speaker 112a, two surround sound speakers 112b and 112c, and two widened front speakers 112d and 112e.

[0083] In this example, each of five input signals 802a-802e is received and filtered by one of five filters 804a-804e. The filtered input signal 802c corresponds to the virtual center speaker 112a and need not be filtered or processed further. The filtered input signals 802a and 802b that correspond to the front virtual speakers 112d and 112e are used to form the output signals 818 and 820. These filtered input signals 802a and 802b are also provided to two forward crossover paths 808a and 808b, each of which includes a filter 810a. Similarly, the filtered input signals 802d and 802e corresponding to the virtual surround sound speakers 112b and 112c are provided to two forward crossover paths 808c and 808d, each of which includes a filter 810b.

[0084] The output signals 818 and 820 are fed through two

feedback crossover paths 816a and 816b that each includes a filter 824. Additional output signals 818 and 820 are then produced by combining various ones of the filtered input signals 802, the signals produced by the forward crossover paths 808, and the signals produced by the feedback crossover paths 816.

[0085] In particular embodiments, the front virtual speakers 112d and 112e can be placed at any suitable location, such as locations having an angle 306 of between 50° to 80°. The virtual center speaker 112a is typically placed at an angle 306 of zero degrees, and the filter 804c has a frequency response with the same amplitude as $\frac{H_i(0°)}{H_i(\theta)}$. A forward crossover path need not be provided for the virtual center speaker 112a because the filter in the forward crossover path would have a response of $\frac{H_c(\phi)}{H_i(\phi)}$ (which equals one) without any delay. As a result, a forward crossover path is not needed, although one could still be provided if desired.

[0086] The frequency response P_F of the filters 804a and 804b may equal the amplitude of $\frac{H_i(\omega)}{H_i(\theta)}$, where ω is the azimuth of the front virtual speakers 112d and 112e. Low-pass filters could be used for filters 810a to approximate $\frac{H_c(\omega)}{H_i(\omega)}$. For non-individualized

design, the azimuth could be assumed to equal 70°, and the angle 304 could be assumed to equal 20°. In this example, a filter with a response shown in FIGURE 8B can be used for filters 804a and 804b, and a first order IIR low-pass filter with a cut-off frequency at 1000 Hz and an attenuation of 3 decibels can be used for filters 810a. The amplitude of the frequency response P_C for filter 804c may equal the amplitude of $\frac{H_i(0°)}{H_i(\theta)}$. A non-individualized design for filter 804c could be a filter with a response shown in FIGURE 8C.

[0087] The various virtualizers 210 shown in FIGURES 4-6, 7A, and 8A and the various frequency responses shown in FIGURES 7B, 8B, and 8C are for illustration only. Other designs or arrangements for the virtualizer 210 could be used without departing from the scope of this disclosure. Also, the different embodiments of the virtualizer 210 shown in the figures could be used in the same audio/video device 102. For example, the virtualizer 210 could be implemented using a DSP that can be reconfigured depending on the mode selected by a listener 108. This may allow, for example, the listener 108 to select a suitable operating mode when the audio/video device 102 is used in different circumstances.

[0088] FIGURE 9 illustrates an example method 900 for rendering audio information to virtualize one or more speakers 112 according

to one embodiment of this disclosure. The method 900 is described with respect to the virtualizer 210 of FIGURE 8A operating in the audio/video device 102 of FIGURE 2B. Other virtualizers or devices could use the method 900 without departing from the scope of this disclosure.

[0089] The audio processor 204 configures the virtualizer 210 at step 902. This may include, for example, the controller 212 of the audio processor 204 using the parameters stored in the memory 206 to configure the filter responses and delay lines in the virtualizer 210.

[0090] The audio processor 204 receives input signals for one or more audio channels at step 904. This may include, for example, the virtualizer 210 receiving five channels from an audio decoder 250, where the channels are supported by the 5.1 rendering standard.

[0091] The audio processor 204 filters one or more of the input signals at step 906. This may include, for example, the virtualizer 210 filtering one, some, or all of the input signals.

[0092] The audio processor 204 provides one or more of the filtered signals to one or more forward crossover paths at step 908. This may include, for example, the virtualizer 210 providing a filtered input signal for a virtual center speaker 112a, a virtual surround sound speaker 112b or 112c, or a virtual forward

speaker 112d or 112e to a forward crossover path. This may also include the virtualizer 210 providing one, some, or all of the filtered input signals to one or more forward crossover paths.

[0093] The audio processor 204 provides one or more previously generated output signals to one or more feedback crossover paths at step 910. This may include, for example, the virtualizer 210 providing one or more previously produced output signals to one or more feedback crossover paths. This may also include the feedback crossover paths generating one or more out-of-phase signals, which are used to reduce or eliminate crosstalk.

[0094] The audio processor 204 produces one or more additional output signals at step 912. This may include, for example, the virtualizer 210 using one or more combiners to combine various ones of the original input signals, the filtered input signals, the signals produced by one or more of the forward crossover paths, and the signals produced by one or more feedback cross over paths.

[0095] Although FIGURE 9 illustrates one example of a method 900 for rendering audio information to virtualize one or more speakers 112, various changes may be made to FIGURE 9. For example, while FIGURE 9 shows various steps occurring sequentially, various steps could also be performed concurrently by the audio processor 204. As a particular example, steps 906-912 could operate concurrently when the audio processor 204 receives input audio signals.

[0096] This disclosure has described the virtualization of one or more virtual speakers 112 in a two-speaker system 100. However, the same or similar principles can be used to virtualize any number of virtual speakers 112 in a system having any number of physical speakers.

[0097] While this disclosure has described certain embodiments and generally associated methods, alterations and permutations of these embodiments and methods will be apparent to those skilled in the art. Accordingly, the above description of example embodiments does not define or constrain this disclosure. Other changes, substitutions, and alterations are also possible without departing from the spirit and scope of this disclosure, as defined by the following claims.